

Approximate Optimal Control of the Compass Gait on Rough Terrain

Katie Byl and Russ Tedrake

Abstract—In this paper, we explore the capabilities of actuated models of the compass gait walker on rough terrain. We solve for the optimal high-level feedback policy to negotiate a perfectly known but qualitatively complex terrain, using a fixed low-level controller which selects a high-level action once-per-step. We also demonstrate that a one-step time horizon control strategy using the same low-level controller can provide performance which is surprisingly comparable to that of the infinite time horizon optimal policy. The model presented here uses a torque at the hip and an axially-directed impulsive toe-off applied just before each ground collision. Our results provide compelling evidence that actuated robots based on passive dynamic principles (e.g. no ankle torque) should inherently be capable of walking on significantly rough terrain.

I. INTRODUCTION

For legged robotic locomotion to become practical outside of a controlled, laboratory setting, robots will need robust, dynamic control strategies which can cope with the variability in environmental conditions expected in day-to-day operation. Such robots will encounter differences in terrain profiles and in characteristics such as contact friction and coefficient of restitution, and they will be subject to intermittent external forces.

Passive dynamic principles have inspired a promising avenue of research in legged robotics ([12], [7], [11], [4], [14], [16], [6]). Purely passive walkers, which require no actuation or active control, exhibit stable, limit-cycle walking down a gentle slope. Unfortunately, the coupling of the stance- and swing-leg dynamics in a passive walker requires a narrow set of acceptable initial conditions (about some nominal, post-collision *fixed point* in state space) for any particular step to be completed successfully. As a result, passive walkers have notoriously fragile basins of attraction; minor perturbations, such as subtle variations in terrain, tend to knock them away from the fixed point and to cause them to fall. Work has been done toward the problem of walking on varying terrain, but in general, the methodologies developed to date have some limitations. These methods include a passivity-passed approach [14], which requires a fully-actuated ankle torque rather than the passive ground contact found in our compass gait model, and a manifold control method [6], which is currently limited to gradual variations in slope. Additionally, the sample-based hybrid-zero dynamic (HZD) approach [16] is very promising as a low-level solution for efficient and robust control, but it still does not solve the high-level problem of setting the low-level control at each step to walk on more complex terrain profiles.

K. Byl and R. Tedrake are with the MIT Computer Science and Artificial Intelligence Lab, Cambridge, MA 02139, USA. {katiebyl, russt}@mit.edu. This work was supported by MIT.

In this paper, we present highlights of our results in the study of a compass gait (CG) walker model with torque actuation at the hip and an axially-directed source of impulse at each toe; the model is described in more detail in Section II. Our goal is to assess the capabilities of this machine on known (but complex) upcoming terrain; no attempt has been made to consider energy efficiency.

II. METHODOLOGY

To solve for the approximate optimal control solution for the compass gait walker on rough terrain, we discretized the dynamics, creating a mesh of post-collision states. We then used the value iteration algorithm [15] to find an optimal step-to-step feedback policy for the discretized system. We tested two low-level control strategies for defining the state-to-state dynamics. The first method was a Proportional-Derivative (PD) controller regulating the inter-leg angle, α . The second method was an impulsive toe-off, applied just before the next ground collision. We tested each method alone and also tested both together. Details on the simulations are given below in the remainder of this section.

A. The Compass Gait Model

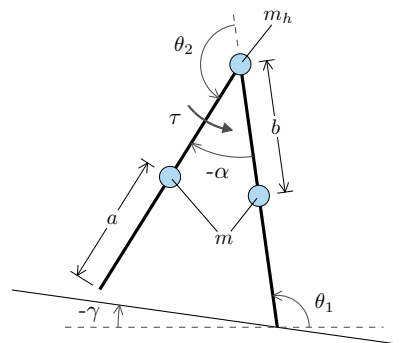


Fig. 1. The compass gait piped model. Simulations in this paper used the following parameter values: $m = 2\text{kg}$, $m_h = 2\text{kg}$, $a = .5\text{m}$, $b = .5\text{m}$

The basic compass gait model, introduced by [7], is used in all simulations and depicted in Figure 1. It consists of two rigid legs: a stance leg (at absolute angle θ_1) and a swing leg (at angle θ_2 , relative to θ_1). The legs meet at a linkage referred as the “hip”, which has a pure torque source. The contact point with the ground is referred to throughout as the “toe” and it is a zero-torque, frictionless joint. The model has a distributed inertia, represented by mass m_h at the hip and a mass m along each leg, as illustrated. Ground collisions are assumed to be instantaneous and perfectly inelastic, and the labels of stance and swing leg are interchanged at each step. Parameters for the model used throughout this paper

are given in the caption of Figure 1, and the direction of walking for all simulations is from the left to the right.

To address the practical concerns of defining and detecting ground collision, we assume that the toe of the swing leg is immediately retracted once a ground collision occurs and remains so until the stance leg reaches a vertical position, at which time it extends instantaneously. We also include an ideal impulsive source at each toe which we assume can deliver an impulse of prescribed magnitude axially at the stance leg contact immediately preceding an upcoming collision between the swing leg and the ground. Inclusion of an impulse source is inspired by the work of [10], who illustrate that this is an efficient method for imparting energy to the system to make up for energy lost in ground collisions. They use the simplified version of the compass gait (i.e. all mass concentrated at the hip; infinitesimal mass at each toe).

B. Optimal Control Using Value Iteration

For each of the three low-level control strategies, we discretized the dynamic system and found the optimal high-level policy which maximized the distance traveled over time before falling. To determine the *full* capabilities of the walker, one would need to discretize over time and to select a torque magnitude at each dt . We chose to discretize on a step-to-step basis instead, primarily to reduce the size of the state space (from 5 to 4 dimensions) to keep the problem computationally tractable, and we were pleased to find that such a strategy still produces impressive results in our simulations.

The post-collision state of the walker is represented in most plots using four *meshing* state variables, X_{1m} through X_{4m} , which are defined below.

- X_{1m} : absolute x coordinate of stance leg on terrain
- X_{2m} : relative x coordinate of swing leg from stance leg
- X_{3m} : absolute angular velocity of stance leg, $\dot{\theta}_1$
- X_{4m} : relative angular velocity of swing leg, $\dot{\theta}_2$

The table below gives the number of elements and range spanned for each meshing state. Mesh elements are spaced evenly in each dimension.

State	# elements	min. value	max value	units
X_{m1}	140	0	7	(m)
X_{m2}	15	-0.85	-0.15	(m)
X_{m3}	14	-3	-0.4	(rad/s)
X_{m4}	14	-0.1	5.1	(rad/s)

The value function and feedback policy between the discretization points was smoothed using barycentric interpolation [13]. To maximize the distance traveled, the cost to optimize was defined as the negative of the distance traveled when a successful step was taken, and as 0 at each step for all walkers which have entered an absorbing, “fallen” state. The value function and feedback policy between the discretization points was smoothed using barycentric interpolation [13]. Any post-collision states falling outside the defined mesh range were automatically binned into the absorbing failure (fallen) state.

C. Rough Terrain Model

We attempted to solve the optimal control problem for each controller on a variety of terrain profiles, some of which include discontinuities in addition to changes in slope. Examples are shown in Figures 2, 4, 6 and 8.

To test the performance limits of each control strategy analyzed, each terrain was scaled to make its features more dramatic until the value iteration failed to converge on a stable walking policy. Each terrain consists of a particular profile which repeats every 7 meters, as shown in Figure 2. This allows the value iteration to converge on a fixed policy using a finite representation for the terrain and a finite number of iteration steps. A repeating terrain may also be thought to represent the terrain of a compass gait walker circling endlessly on a fixed boom; we intend to experiment with the control strategies described in this paper on a real robot mounted to a such a boom, as depicted in Figure 10 in Section VI.

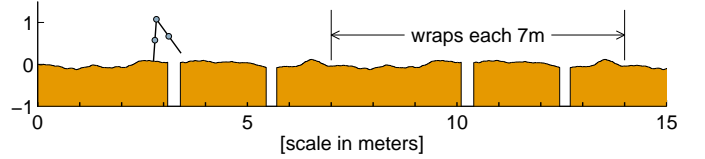


Fig. 2. An example of terrain. Tested terrain profiles always repeat every 7 meters, allowing the numerical mesh to “wrap”.

D. Hierarchical Controller Design

Our simulations investigate two simple low-level control options, and optimal policies are found for each of three combinations: using each option individually, and using them together. One low-level option is Proportional-Derivative (PD) control of a desired inter-leg angle, and the other is impulsive stance-foot toe-off immediately preceding the ground collision for each new step. The high-level control action selected at each step in the value iteration correspondingly consists of one or both of the following: (1) a desired inter-leg angle, α_{des} , for the PD controller and (2) the magnitude of the impulse.

The primary purpose of the PD controller is to regulate of the step length of the walker, which in turn selects upcoming foot placement on the terrain. However, the dynamics are inherently coupled and the controller also affects the entire dynamic state of the walker. The main goal in employing the impulsive toe-off action is to compensate for the energy which is lost at each ground collision. This in turn should allow the walker to take larger steps than would otherwise be possible, since more energy is (of course) lost for larger step angles [3].

1) *PD Control of Inter-leg Angle*: The low-level PD controller regulates the inter-leg angle, α , defined as:

$$\alpha = \theta_2 - \pi \quad (1)$$

Our PD controller was designed by hand, using simulations of the compass gait walker on flat terrain and observing

performance given intermittent, impulsive disturbances. It acts only during the latter part of each step, after the stance leg has passed vertical. Activating the torque too early in the gait tends to cause the walker to fall over backward before the stance leg has successfully passed beyond the vertical point in its trajectory. The hip torque, τ , is defined by Equation 2:

$$\tau = \begin{cases} K_p(\alpha_{des} - \alpha) + K_d(0 - \dot{\alpha}) & , \text{ if } \theta_1 < \frac{\pi}{2} \\ 0 & , \text{ otherwise} \end{cases} \quad (2)$$

where $K_p = 100$ and $K_d = 10$.

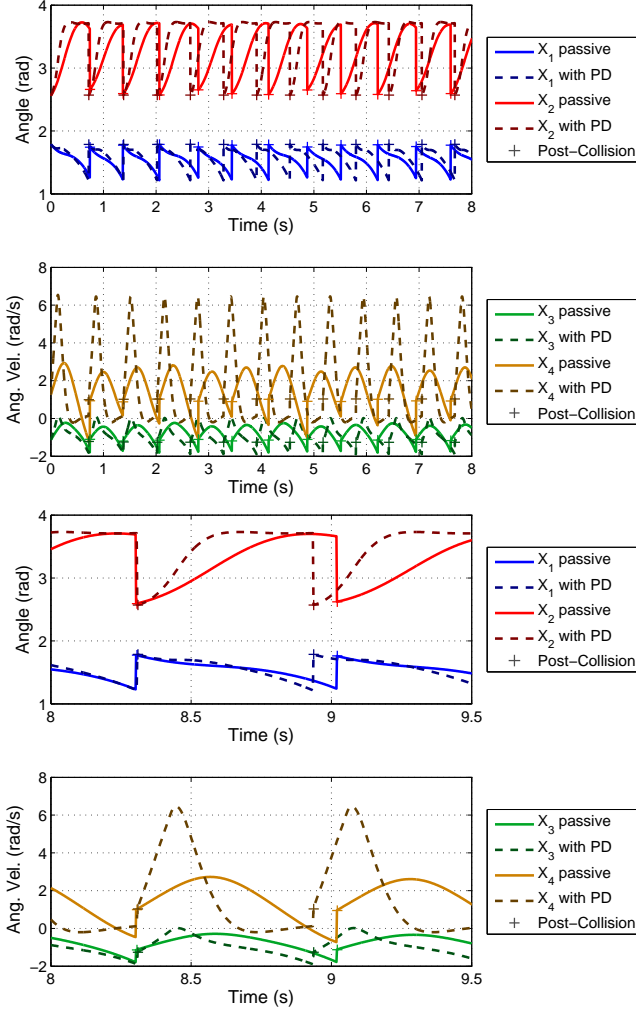


Fig. 3. Comparison of passive and PD-controlled walker on a constant 4° slope. Although the peak swing-leg velocity each cycle is over 2 times greater for the PD walker, the overall speed in walking is only about 9% faster, since the dynamics are still dominated by the passive, inverted-pendulum motion of the stance leg.

2) *Impulsive Toe-Off Immediately Before Impact:* In this control strategy, a particular magnitude impulse is applied axially from the stance leg downward, toward ground. It is well-known that applying the impulse immediately before collision is an efficient method for imparting energy to the compass gait walker [10]. We assume the impulse is executed perfectly and instantaneously.

The high-level feedback policy is evaluated at each *post-collision* state, but we desire that a commanded *pre-collision*

impulse action affect the current step. Therefore, to model the transition, the post-collision state is simply “rewound” to its corresponding pre-collision state, using the well-known equations for conservation of angular momentum during the inelastic collision [7]. Then, the contribution of the impulse ($m \times v$) is added as appropriate at the toe of the stance leg, and the equations for angular momentum (now going from pre- to post-collision) are applied once again. Note that the velocity of the swing leg must be tested after the impulse is applied to ensure that the velocity of the swing leg toe is still directed toward the ground. A large enough impulse could theoretically send the entire walker airborne! However, this has not been a practical concern for the magnitudes we have tested in our simulations.

Our initial tests using value iteration allowed for the selection of one of 21 values ranging from 0 to 2 (kg-m/s). Because the cost function described below rewards taking as large a step as possible, the optimal impulse selected was almost always the maximal value possible. To simplify the controller and to reduce the run-time for value iteration (by a factor of 21), we eventually set the impulse value to a constant magnitude of 2 at every step.

E. Control Using a Single-Step Time Horizon

Our main purpose in this paper is to demonstrate the theoretical capabilities inherent in the underactuated walker as terrain becomes more extreme, and so we provide perfect long-horizon knowledge of the terrain to decouple this factor from the issue of what is dynamically possible, given the underactuation. However, to further explore the theoretical performance of the compass gait model on rough terrain, we also tested a one-step control strategy. At each step, a new action is selected by simulating a nominal value for α_{des} for one step, assuming we have access to a perfect model of the one-step dynamics. If the tested value results in a successful (non-fallen) upcoming step, it is selected as the desired action. If not, we continue to search inter-leg angles in a discrete, prescribed sequence until a successful candidate angle is found.

The one-step control policy we implement is extremely primitive, yet it has proven surprisingly effective. This result provides some evidence that use of a very limited look-ahead (e.g. 1 to 3 steps) may provide near-optimal performance. Correspondingly, it follows that it may also be possible to obtain near-optimal results on real-world legged robots having only a short-sighted knowledge of upcoming terrain. We anticipated this result after our previous work, where we noted a surprisingly fast mixing dynamics of the metastable limit cycle generated by a passive compass gait walker on statistically rough terrain [2].

III. RESULTS

This section highlights our simulation results for each of the three control strategies described in Section II. Summarized briefly, the PD controller is effective in regulating foot placement on terrain; however, it allows for a disappointingly limited variations in step width or height. The toe-impulse

action is *intended* to enable the walker to negotiate more extreme terrain by pumping additional energy into the system efficiently, but performance of the impulse control alone is unfortunately quite fragile. Combining a constant magnitude toe-off with the PD controller provides a significantly better design that either component demonstrates alone, allowing for significantly greater variations in both step length and height during continuous walking.

A. Performance Using Only PD Control

Figures 4 and 5 show examples of terrain which our simulation successfully negotiates using only the PD controller. The performance is good, but it is importantly limited by the inherent inefficiency of adding energy through hip torque, alone. A single step which is too wide or too steeply uphill results in rapid failure, as the compass gait simply cannot complete a subsequent step after a large loss in energy.

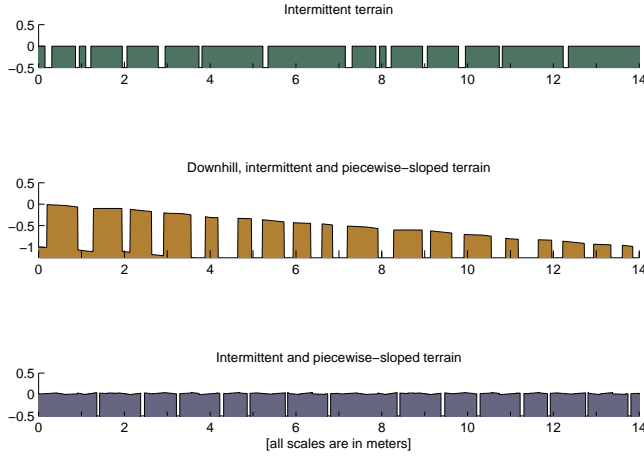


Fig. 4. Examples of terrain which were successfully negotiated using PD control alone. Terrain at top is flat except for a set of bottomless no-go gaps. The middle and bottom examples consist of piecewise slopes. Figure 5 shows detail from the bottom plot.

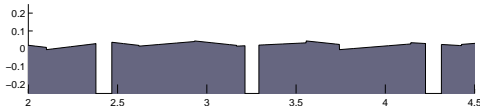


Fig. 5. A close-up of the bottommost terrain in Figure 4.

Several features can be noted about both the performance capabilities of the walker and the nature of the resulting optimal control policy. Qualitatively, the walker is able to perform much larger steps when the terrain is sloped downhill, as one might naturally expect. Also, the more extreme (“difficult”) the terrain was, the more likely the optimal policy was to select the same, repeated set of footholds over a particular portion of the terrain, as illustrated in Figure 6.

The terrains shown throughout the paper are what the authors consider to be the most impressive examples for which successful optimal policies could be found; on more extreme terrain, there were no solutions which resulted in continuous walking. That noted, it is interesting to compare

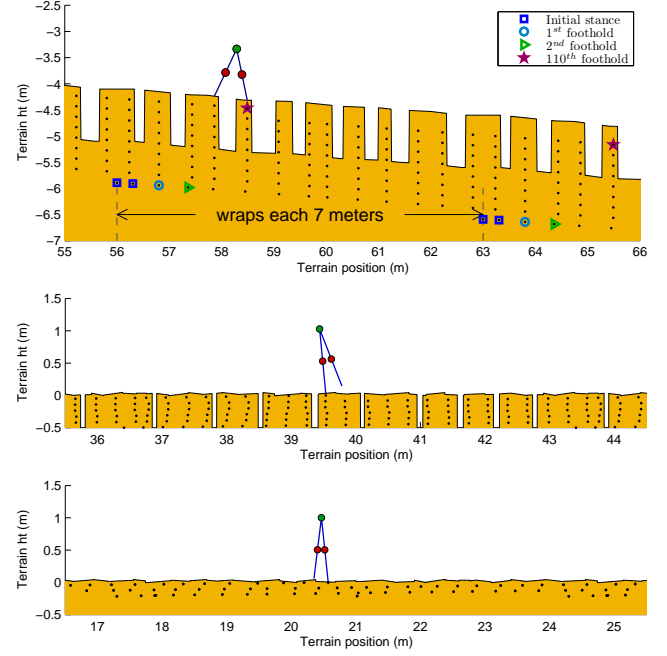


Fig. 6. The optimal footholds on extreme terrain quickly converge to a fixed pattern (top). On easier terrain, no fixed pattern emerges, because the cost function always rewards taking the largest steps possible (bottom). In intermediate terrain, we see an intermediate level of organization in the pattern of footholds (middle). The x locations of footholds taken are plotted as dots here and in Figure 8, with the earliest steps plotted lowest and with dots repeating as the terrain repeats, every 7 meters.

the features of the three types of intermittent terrain in Figures 4. On the flat (topmost) terrain, gap width was about 11 cm on average, up to a maximum of 18 cm; for the piecewise sloped terrain (bottom), the gaps needed to be somewhat smaller: 8 cm on average up to a maximum of about 10 cm. The downhill terrain (shown in the middle) has an average grade of 4° and correspondingly permits dramatically larger steps: 30 cm on average, with the larger step being about 47 cm – or nearly half the leg length.

The overall variation in terrain height away from a constant-slope terrain is not significantly different for the downhill terrain with large steps as compared with the piecewise sloped terrain. Excluding the part of the terrain with intermittent gaps, the piecewise terrain has a standard deviation (SD) of 1.07 cm, and a max-to-min height difference of 5.15 cm. Subtracting out the net slope of 4° , the “downhill” terrain height has a standard deviation of 1.26 cm and max-to-min height different of 6.15 cm. Step width on this terrain *does* vary significantly (from about 28 cm to 63 cm), but it cannot maintain a gait with continuously small footsteps without tripping forward and falling after a few steps. Torque-only actuation is often insufficient to completely regulate energy.

B. Performance Using Only Toe-Off Control

It was hoped that an impulsive toe-off controller might produce better results than the PD controller. However, the toe-off alone results in a surprisingly fragile walking cycle

which requires a fairly precise initial condition and allows for almost no variations in terrain height. We let a walker under impulse-only control reach a steady limit cycle on terrain which is completely flat (0° slope), using a constant impulse value of $m \times v = 2$ (kg-m/s). Then we introduce slight, rolling bumps into the terrain of about 1/12 the magnitude of those seen on the terrain in Figures 8.

After reaching the mild bumps on the terrain, the “toe-off only” walker takes one step with a resulting downward slope of -7° and then “trips up” while attempting a subsequent step at $+35^\circ$. The performance is quite poor, even if one allows for the selection of the magnitude of the impulse. Without more direct regulation of the upcoming step length, the walker quickly falls on even trivial terrain.

Figure 7 shows the dynamic states (angles and angular velocities) of the walker during the same trial described above. These data are presented primarily to allow quantitative and qualitative comparison with the dynamic states during the trials using both PD and impulse control, shown in Figure 9. With the torque actuation from the PD controller, the steps are generally both much larger (X_2 is larger) and faster (there are more steps per second, and X_3 and X_4 are both larger when hip torque is employed). Note that two of the dynamic states plotted in these figures differ from the meshing states defined in Section II-B: $X_1 = \theta_1$ and $X_2 = \theta_2$, while $X_4 = X_{m4} = \dot{\theta}_1$ and $X_3 = X_{m3} = \dot{\theta}_2$.

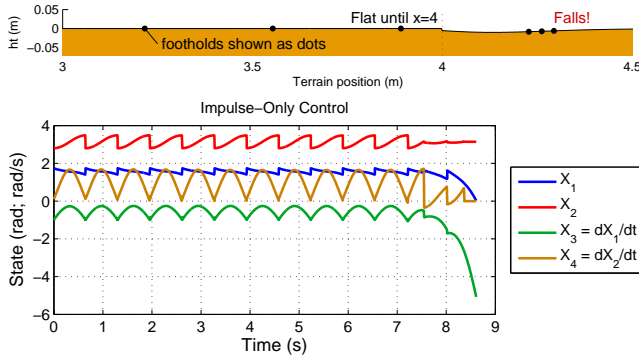


Fig. 7. Dynamic states using impulsive toe-off control, only.

C. Performance Using Both PD Control and Toe-Off

Our most impressive performance by far is obtained by combining the low-level PD controller with the an impulse of magnitude 2 (kg-m/s) preceding every ground collision. The optimal control policy converges to a successful policy for continuous walking on the terrain shown in Figure 8, which has a standard deviation (σ) of 4.7 cm and a max-to-min difference in height of 19.1 cm. The step-to-step heights between footholds actually achieved along this terrain have σ of 3.3 cm, and σ of the inter-leg angle at each step is about 5.9° .

The dual-mode control strategy was so powerful that we were also able to obtain long trials of continuous walking (over 60 seconds in duration) using the simple, one-step control method described in Section II-E. The bottom of

Figure 8 shows footholds taken during one such trial using the one-step planner. Here, the standard deviation in step-to-step height is 5.0 cm, and the SD in α is 8.8° . Finally, Figure 9 depicts the dynamic states for both the optimal and one-step control policies. Note here that the selected set-point for α ($X_2 = \alpha + \pi$) has very little variation in the one-step planner. Deviations in the set point of α are occasionally required for success, but not if a nominal first-tested value (of 28°) works. It is intriguing that this low-level controller works surprisingly well even given such a clearly sub-optimal high-level policy. This leads us to suspect that it is quite probable that a more intelligently-design short-sighted planner might perform nearly as well the optimal (far-sighted) one.

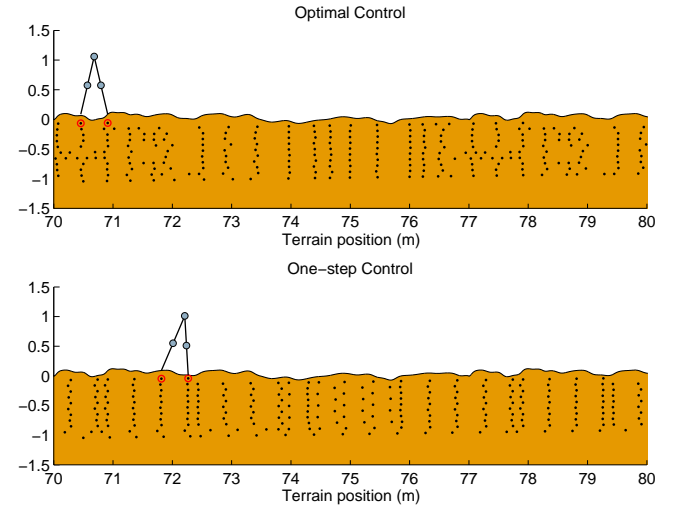


Fig. 8. Footsteps taken during a 60-second trial using the optimal control policy from value iteration (top) and using a one-step time horizon (bottom).

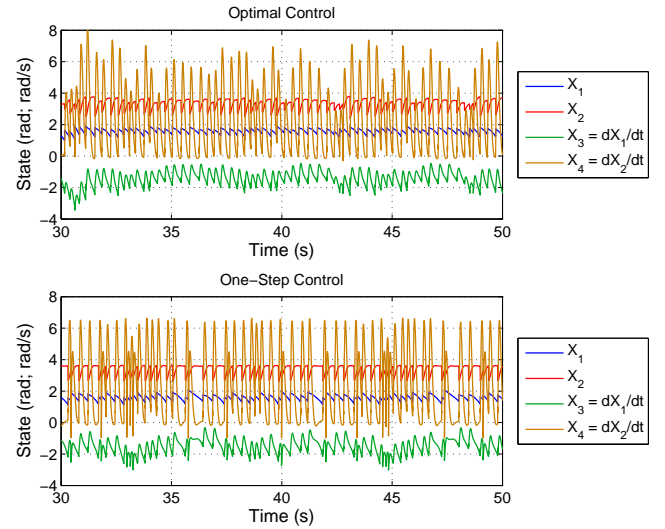


Fig. 9. Dynamic states during a using the optimal control policy from value iteration (top) and using the one-step control strategy (bottom).

IV. DISCUSSION

We wish to bring some particular issues to the attention of the reader in this section. First, we note that care must be taken in interpolating the inherently discontinuous policy for walking on rough (most particularly, on intermittent) terrain to select control actions. We discovered a persistent and catastrophic interpolation error which occurred from time to time when evaluating the optimal policy during simulations. For gaps, there are transition points where one must decide to take either a small step or a large step to avoid such regions. We found there were occasionally but consistently states where barycentric interpolation of the policy resulted in a medium-sized step which landed in a no-go zone. Although [13] provide an excellent reference on issues of stochasticity and *discontinuity of value function* in implementing the solution to a dynamic system via value iteration, their solutions can be expensive in practice and do not directly address our particular interpolation issues, which are associated with sharp *discontinuities in our policy*.

The second issue we bring to the reader's attention here is that our solutions do not attempt to exploit the low cost of transport and efficiency which are so characteristic and appealing in particularly well-designed passive dynamic walkers, such as the Cornell biped [5], [9]. Our purpose here was to demonstrate the theoretical performance possible for passive-based walking with an inherently underactuated ankle. Further study can be done using a cost function which takes into account both the distance traveled and the cost of transport.

Finally, we note that the success of a simple, intuitive strategy for control of the Raibert hopper on intermittent terrain in [8] inspired our inclusion of intermittent gaps in some of our own terrain examples. The simple control ideas of the Raibert hopper have since provided a basis for the practical design of impressively robust control for the dynamic quadruped "BigDog" on rough terrain [1]. We conjecture that the future for dynamic legged bipedal robots may follow a similar path, using lessons learned from the compass gait model as a basis for impressive controllers on advanced robots based on passive dynamic principles.

V. CONCLUSIONS

The compass gait model can successfully negotiate an impressive range of rugged terrain using a surprisingly simple control strategy. A simple yet effective low-level control strategy can be obtained by combining a toe-off just before each ground collision with a PD control loop on the desired inter-leg angle. The magnitude of the toe-off can be a constant value for all footsteps, so that the high-level control at each step consists solely of the desired inter-leg angle. The action of the toe-off insures energy is added efficiently at each step, while actively controlling the inter-leg angle tends to regulate the dynamics around a nominal trajectory in state space.

VI. FUTURE WORK

Given the success of our simulations, we plan to implement a similar control strategy on a real compass-gait robot. The robot will be mounted on a boom, providing lateral stability but also introducing some additional, unmodeled dynamics. We are hopeful that the inherent stability demonstrated by the simple low-level control strategy described in this paper can also be demonstrated on the real-world system. We expect that a controller based on the same practical principles will provide a good initial control policy, to be optimized online using model-free gradient-based learning methods.



Fig. 10. Compass gait robot posed on rough terrain.

REFERENCES

- [1] M. Buehler, R. Playter, and M. Raibert. Robots step outside. *Int. Symp. Adaptive Motion of Animals and Machines (AMAM)*, Ilmenau, Germany, pages 1–4, Sep 2005.
- [2] Katie Byl and Russ Tedrake. Stability of passive dynamic walking on uneven terrain. In Art Kuo, editor, *Proceedings of Dynamic Walking 2006*, May 2006.
- [3] Michael J. Coleman and Andy Ruina. Motions of a rimless spoked wheel: a simple 2d systems with impacts. August 2002.
- [4] Steven H. Collins, Andy Ruina, Russ Tedrake, and Martijn Wisse. Efficient bipedal robots based on passive-dynamic walkers. *Science*, 307:1082–1085, February 18 2005.
- [5] Steven Hartley Collins and Andy Ruina. A bipedal walking robot with efficient and human-like gait. *Proc. IEEE International Conference on Robotics and Automation, Barcelona, Spain*, Apr 2005.
- [6] Tom Erez and William D. Smart. Bipedal walking on rough terrain using manifold control. pages 1539–1544, 2007.
- [7] A. Goswami, B. Espiau, and A. Keramane. Limit cycles and their stability in a passive bipedal gait. pages 246 – 251. *IEEE International Conference on Robotics and Automation (ICRA)*, 1996.
- [8] Jessica Hodgins and Marc Raibert. Adjusting step length for rough terrain locomotion. *IEEE Transactions on Robotics and Automation*, 7(3):289–298, June 1991.
- [9] A. D. Kuo. Choosing your steps carefully. *Robotics and Automation Magazine, IEEE*, 14(2):18–29, 2007.
- [10] A D Kuo, J M Donelan, and A Ruina. Energetic consequences of walking like an inverted pendulum: Step-to-step transitions. *Exerc. Sport Sci. Rev.*, 33(2):88–97, 2005.
- [11] Arthur D. Kuo. Energetics of actively powered locomotion using the simplest walking model. *Journal of Biomechanical Engineering*, 124:113–120, 2002.
- [12] Tad McGeer. Passive dynamic walking. *International Journal of Robotics Research*, 9(2):62–82, April 1990.
- [13] Remi Munos and Andrew Moore. Variable resolution discretization in optimal control. *Machine Learning*, 49(2/3):291–323, November/December 2002.
- [14] Mark W. Spong. The passivity paradigm in bipedal locomotion. In *Proceedings of the International Conference on Climbing and Walking Robots (CLAWAR)*, September 2004.
- [15] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.
- [16] E. R. Westervelt, B. Morris, and K. D. Farrell. Analysis results and tools for the control of planar bipedal gaits using hybrid zero dynamics. *Autonomous Robots*, 23:131–145, Jul 2007.